

Klasifikasi Penjualan Produk Menggunakan Algoritma Naive Bayes pada Konter HP Bayu Cell

¹Pradani Ayu Widya Purnama, ²Teri Ade Putra
^{1,2}Universitas Putra Indonesia “YPTK” Padang
Padang, Indonesia

pradaniwid@gmail.com

*Penulis Korespondensi

Diajukan : 29/12/2023
Diterima : 08/01/2024
Dipublikasi : 12/01/2024

ABSTRAK

Data Mining merupakan proses iteratif dan interaktif untuk menemukan pola atau model baru yang dapat digeneralisasi untuk masa yang akan datang, bermanfaat dan dapat dimengerti dalam suatu database yang sangat besar (massive database). Pada Konter HP Bayu Cell mempunyai masalah dalam menentukan produk yang laris atau tidak laris terjual pada konter tersebut. Dengan demikian untuk adanya suatu sistem yang dapat menunjang Konter HP Bayu Cell dalam mengambil keputusan secara cepat dan tepat agar tidak terjadinya penumpukan produk. Untuk itu pada penelitian ini penulis menggunakan Algoritma Data Mining Metode Algoritma Naive Bayes, Pada metode Naive Bayes untuk ukuran peluang suatu peristiwa yang terjadi berdasarkan peristiwa lain yang telah (dengan asumsi, praduga, pernyataan, atau terbukti) terjadi. Pada data training yang dilakukan dengan produk handphone REDMI NOTE 9 menggunakan algoritma Naive Bayes dimana menghasilkan klasifikasi masuk ke kategori “**Laris**” dengan nilai **0,005211**. Oleh karena itu pada konter HP Bayu sudah dapat menghasilkan hasil yang diharapkan dan dapat membantu dalam mengetahui mana saja produk handphone yang termasuk kategori laris dan tidak laris.

Kata Kunci: Data Mining, Naive Bayes, Penjualan

I. PENDAHULUAN

Data mining digunakan untuk meningkatkan keuntungan ataupun membantu dalam klasifikasi dalam penjualan produk. Salah satu penggunaan data mining dengan metode Naive Bayes dalam data penjualan umumnya adalah untuk mengetahui minat dan ketertarikan calon pembeli terhadap produk yang tersedia dengan memprediksi probabilitas keanggotaan suatu class. Informasi ini dapat digunakan untuk mendukung strategi pemasaran agar lebih efektif dan efisien. Dengan informasi ini suatu perusahaan dapat mengetahui tingkat ketertarikan pembeli terhadap suatu produk yang terdapat di perusahaan. Sehingga perusahaan dapat mengetahui serta menentukan target atau sasaran pasar dengan lebih rinci (Kursini, et al. 2009).

Algoritma Naive Bayes adalah classifier probabilistik sederhana yang menghitung seperangkat probabilitas dengan menghitung frekuensi dan kombinasi nilai pada kumpulan data yang diberikan. Algoritma ini menggunakan teorema Bayes dan mengasumsikan semua atribut menjadi independen mengingat nilai variabel kelas ini kondisional dengan asumsi bahwa kemerdekaan jarang berlaku pada aplikasi dunia, maka karakteristik diasumsikan sebagai Naif namun algoritma cenderung berkinerja baik dan dapat belajar dengan cepat dalam berbagai masalah klasifikasi. Naive Bayes adalah suatu klasifikasi kemungkinan sederhana yang dapat

menghitung seluruh kemungkinan dengan menggabungkan sejumlah kombinasi dan frekuensi suatu nilai dari basis data yang didapatkan (Rachman, et al. 2021).

Menurut Indah Purnamasari, 2018 dalam penelitiannya yang berjudul Klasifikasi Pelanggan Produk Indihome Menggunakan Naive Bayes Classifier dengan Seleksi Fitur Algoritma Genetik didapatkan hasil perhitungan tersebut diketahui nilai $P(H|3P)$ lebih besar dari pada nilai $P(H|Decline)$, sehingga dapat disimpulkan bahwa untuk kasus tersebut masuk kedalam klasifikasi 3P.

Berdasarkan penelitian yang dilakukan Permadi, V. A. 2020. Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran Di Singapura hasil pengujian yang dilakukan, pengklasifikasian sentimen dengan algoritma naive bayes memberikan nilai precision sebesar 73,02%, recall sebesar 74%, serta akurasi sebesar 73,33%.

Berdasarkan penelitian yang dilakukan Rachman, et al. 2021 dengan judul Klasifikasi Algoritma Naive Bayes dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM, menghasilkan nilai Accuracy 81.81%, Precision 66.66% dan Recall 100% dengan hasil AUC 0.800.

Penelitian selanjutnya oleh Rayuwati, et al. 2022 dengan judul Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebab Covid dengan memprediksi tingkat penyebaran covid-19 di wilayah kabupaten aceh tengah dengan tingkat akurasi sebesar 60%.

Naive Bayes telah digunakan sebagai pengklasifikasi yang efektif bertahun-tahun. Naive Bayes mudah untuk dibangun karena strukturnya diberi atribut bersifat independent untuk memecahkan masalah. Oleh karena sifat independent ini maka performa Naive Bayes Classifier menjadi tidak sempurna, untuk meningkatkan kinerja Naive Bayes perlu dioptimasi dengan seleksi fitur. Pada penelitian ini Klasifikasi untuk melihat penjualan produk pada Konter HP Bayu Cell Menggunakan Algoritma Naive Bayes.

II. STUDI LITERATUR

Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau mencari dari data yang terdapat pada basis data. Data mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut Knowledge Discovery Databases (KDD). Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifikasi ada dua pekerjaan utama yang dilakukan, yaitu Pembangunan model sebagai prototype untuk disimpan sebagai memori dan Penggunaan model tersebut untuk melakukan pengenalan/klasifikasi/prediksi pada suatu objek data lain agar diketahui dikelas mana objek data tersebut dalam model yang sudah disimpannya (Kursini, et al. 2009).

Naive Bayes

Naive Bayes adalah suatu klasifikasi kemungkinan sederhana yang dapat menghitung seluruh kemungkinan dengan menggabungkan sejumlah kombinasi dan frekuensi suatu nilai dari basis data yang didapatkan. Suatu algoritma memanfaatkan teorema bayes dan memperkirakan seluruh atribut yang bebas dan saling lepas yang dapat diberikan oleh suatu nilai pada kelas variable (Rachman, et al. 2021).

naive bayes adalah algoritma pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Algoritma naive bayes terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar

Helmayanti, et al. (2023).

Teorema Bayes memiliki bentuk umum sebagai berikut

$$P(H|X) = \frac{P(X|H) P(H)}{P(X)} \dots\dots\dots (1)$$

Di mana :

X = Kelas data yang belum diketahui

H = Hipotesa data X adalah kelas spesifik

P(H|X) =Kemungkinan Hipotesa H berdasarkan keadaan X (posteriori prob.)

P(H) = Kemungkinan Hipotesa H (prior prob.)

P(X|H) = Kemungkinan X berdasarkan keadaan tersebut

P(X) = Kemungkinan dari X

Naive Bayesian Classifier mengasumsikan bahwa keberadaan sebab atribut (variabel) tidak ada kaitannya dengan beradaan atribut (variabel) yang lain karena asumsi atribut tidak saling terkait(conditionally independent), ditulis dengan rumus :

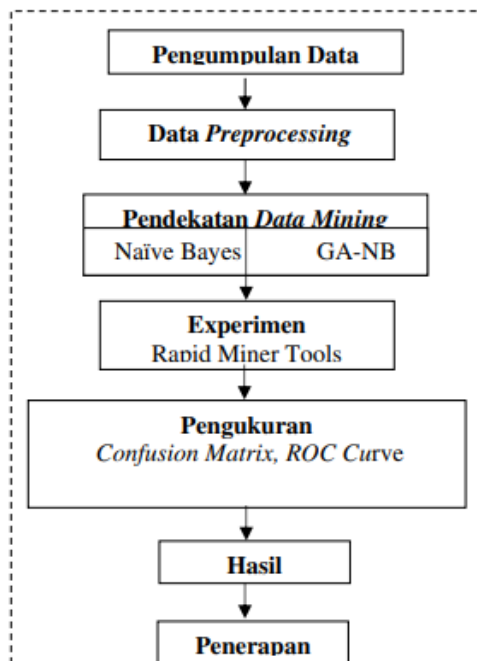
$$P(X|Ci) = \prod_{k=1}^n P(Xk|Ci) \dots\dots\dots (2)$$

Setelah diperoleh hasil dari seluruh data pada setiap clas, maka hasil akhirnya dapat menggunakan rumus :

$$P(X|Ci) = \arg \max P(Xi|Ci) * P(Ci) \dots\dots\dots (3)$$

III. METODE

Kerangka kerja ini merupakan langkah-langkah yang akan dilakukan dalam rangka penyelesaian masalah yang akan dibahas. Berikut adalah kerangka kerja penelitian dapat dilihat pada Gambar 1.



Gambar 1. Kerangka kerja

Berdasarkan kerangka kerja penelitian pada Gambar 1, dapat dilihat bahwa penelitian ini menggunakan data penjualan produk sebagai data input. Penelitian ini menggunakan Algoritma Naive Bayes.

IV. HASIL DAN PEMBAHASAN

Konter HP Bayu Cell pada beberapa bulan terakhir. Berikut merupakan data training dalam sistem klasifikasi produk pada Konter HP Bayu Cell:

Tabel 4. 1 Data Probabilitas

No	Merk Hp	Spesifikasi	Range Harga	Garansi	Kondisi	Klasifikasi
1	REDMI 10C 4/128	Middle End	Kurang Rp. 2.000.000	Resmi	Secound	Tidak Laris
2	REDMI 10 2022 4/64	Low End	Kurang Rp. 2.000.000	Resmi	Baru	Laris
3	REDMI 10 2022 6/128	Low End	Kurang Rp. 2.000.000	Toko	Baru	Laris
4	REDMI 10 5G 4/128	High End	Kurang Rp. 2.000.000	Resmi	Secound	Tidak Laris
5	REDMI 10 5G 6/128	High End	Rp.2.000.001- Rp.3.000.000	Resmi	Baru	Laris
6	REDMI NOTE 11 4/128	High End	Rp.2.000.001- Rp.3.000.000	Toko	Baru	Laris
7	REDMI NOTE 11 6/128	High End	Rp.3.000.001- Rp.4.000.000	Resmi	Secound	Tidak Laris
8	REDMI PAD 6/128	High End	Rp.4.000.000	Toko	Baru	Laris
9	REDMI 11 PRO 6/128	High End	Rp.4.000.000	Toko	Secound	Laris
10	REDMI 11 PRO 8/128	High End	Rp.4.000.001- Rp.5.000.000	Resmi	Baru	Laris
11	REDMI A1 2/32	Low End	Rp. 1.149.000	Resmi	Baru	Tidak Laris
12	REDMI A1 3/32	Middle End	Rp. 1.249.000	Resmi	Baru	Tidak Laris
13	REDMI 9A 2/32	High End	Rp. 1.119.000	Toko	Baru	Laris
14	REDMI 9A 3/32	Low End	Rp. 1.229.000	Resmi	Secound	Laris
15	REDMI 9C 3/32	Low End	Rp. 1.349.000	Toko	Baru	Laris
16	REDMI 9C 4/64	High End	Rp. 1.549.000	Resmi	Baru	Laris
17	REDMI 10A 3/32	Low End	Rp. 1.399.000	Resmi	Baru	Laris
18	REDMI 10A 3/64	High End	Rp. 1.499.000	Resmi	Secound	Tidak Laris
19	REDMI 10A 4/64	High End	Rp.3.000.001- Rp.4.000.000	Resmi	Secound	Tidak Laris
20	REDMI 10C 4/64	High End	Rp.5.000.001- Rp.6.000.000	Resmi	Baru	Tidak Laris

Dalam penelitian ini digunakan dua kelas yaitu laris dan tidak laris. Sebelum mencari nilai mean dan standar deviasi, kita harus mengelompokkan terlebih dahulu jumlah kelas laris dan tidak laris berdasarkan data training yang digunakan.

1. Kelas laris = 12 data.

2. Kelas Tidak Laris = 8 data

Probabilitas laris :

$$P(C) = \frac{12}{20} = 0.6$$

Probabilitas Tidak Laris :

$$P(C) = \frac{8}{20} = 0.4$$

Berikut merupakan tabel probabilitas klasifikasi dalam sistem klasifikasi produk pada Toko Bayu Cell:

Tabel 4. 2 Data Probabilitas

Label	Jumlah Data	Total Data	Probabilitas
Laris	12	20	0.6
Tidak Laris	8		0.4

Setelah mencari nilai probabilitas, selanjutnya kita melakukan perhitungan *posterior* tiap-tiap data training.

$$\text{Rumus} = P(C|X) = \frac{P(X|C)P(C)}{P(X)} \tag{1}$$

Dimana :

X : Data dengan kelas yang belum diketahui

C : Hipotesis data X merupakan suatu kelas spesifik

P (C|X) : Probabilitas hipotesis C berdasar kondisi X (probabilitas posterior)

P(C) : Probabilitas hipotesis C (probabilitas prior)

P(X|C) : Probabilitas X berdasarkan kondisi pada hipotesis

C P(X) : Probabilitas X.

Menghitung nilai peluang Merek Hp

$$= X|(Co))P>Nama)) = X|(Co))$$

$$P>Nama = " REDMI 10 5G" | klasifikasi = "Laris")P>Nama = 1/12) = 0.083.$$

$$P>Nama = " REDMI 10 5G" | klasifikasi = "Tidak Laris")P>Nama = 1/8) = 0.125.$$

Berikut merupakan tabel peluang nama Hp dalam sistem klasifikasi produk pada Konter HP Bayu Cell:

Tabel 4. 3 Data Probabilitas

No	Nama	Jumlah Kejadian		Probabilitas	
		Laris	Tidak Laris	Laris	Tidak Laris
1	REDMI 10 5G	1	1	0.083	0.125
2	REDMI 10C	0	2	0	0.25
3	REDMI 10 2022	2	0	0.166	0
4	REDMI NOTE 11	1	1	0.083	0.125
5	REDMI PAD	1	0	0.083	0
6	REDMI 11 PRO	2	0	0.166	0
7	REDMI A1	0	2	0	0.25
8	REDMI 9A	2	0	0.166	0
9	REDMI 9C	2	0	0.166	0
10	REDMI 10A	1	2	0.83	0.25

Nilai Distribusi Gaussian

Langkah selanjutnya adalah menghitung nilai probabilitas untuk fitur Data Testing. Berikut merupakan data testing dalam sistem klasifikasi produk pada Konter HP Bayu Cell:

Tabel 4. 4 Data Testing

No	Merk Hp	Spesifikasi	Range Harga	Warna	Garansi	Kondisi
1	REDMI NOTE 9	High End	Rp.4.000.000	Stellar Black	Resmi	Baru

Hasil belum diketahui klasifikasinya karena variabel tersebut merupakan hasil prediksi dari data yang dihitung dari setiap variabel. Untuk menghitung nilai distribusi gaussian. Berikut adalah proses menghitung nilai distribusi gaussian.

$$= P(\text{Merk Hp} = \text{"REDMI NOTE 9"} | \text{"Laris"}) * \text{Spesifikasi} = \text{"High End"} | \text{"Laris"} * \text{Warna} = \text{"Stellar Black"} | \text{"Laris"} * \text{Range Harga} = \text{"Kurang Rp.4.000.000"} | \text{"Laris"} * \text{Garansi} = \text{"Resmi"} | \text{"Laris"} * \text{Kondisi} = \text{"Baru"} | \text{"Laris"}) * P(\text{Klasifikasi} = \text{"Laris"})$$

$$= (0.166 * 0.6 * 0.166 * 0.416 * 1.166 * 1.083) * 0.6 = 0,005211$$

dapat disimpulkan bahwa produk handpone REDMI NOTE 9 masuk dalam klasifikasi "Laris".

V. KESIMPULAN

Berdasarkan hasil pengolahan menggunakan algoritma Naïve Bayes menghasilkan nilai untuk REDMI NOTE 9 **0.005211** termasuk klasifikasi **Laris**. Pengolahan data yang dilakukan mengasilkan informasi tentang klasifikasi produk HP termasuk Laris atau Tidak Laris, sehingga pihak konter HP Bayu Cell dapat dengan mudah mencek apa saja produknya yang termasuk Laris atau Tidak untuk melakukan stock barang. Untuk penelitian selanjutnya, bisa untuk menggunakan metode lain yang lebih akurat nilainya. Hasil dari penelitian ini dapat dijadikan sebagai salah satu referensi bagi pihak konter dalam mendukung keputusan untuk melakukan tindakan yang bisa meminimalisir terjadinya penumpukan barang.

VI. REFERENSI

- Helmayanti, et al. (2023). Penerapan Algoritma TF-IDF dan Naive Baes untuk Analisis Sentimen Berbasis Aspek Ulasan Aplikasi Flip Google Pla Store. *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*. Vol.4 No.3, September 2023.
- Indah Purnamasari. (2018). Klasifikasi Pelanggan Produk Indihome Menggunakan Naive Bayes Classifier Denagan Seleksi Fitur Algoritma Genetik. *Jurnal Teknik Komputer*. Vol.4 No.1.
- Nanda, et al. 2022. Klasifikasi Berita Menggunakan Metode Support Vektor Machine. *Jurnal Nasional* Vol5 No.2 April 20233.
- Rachman, et al.. (2021). Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM. *Jurnal Informatika*. Vol.8 No.2, Halaman 111-122.

-
- Rayuwati, et al. 2022. Implementasi Algoritma Naive ayes Untuk Memprediksi Tingkat Penyebaran Covid. Jurnal Jusrritek Vol 1 No.1 April, Hal 38-46.
- Kursini, et al. 2009. Algoritma Data Mining. Yogyakarta : Penerbit Andi.
- Permadi, V. A. 2020. Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran Di Singapura. Buana Informatika, Vol.11, No.2:141–151.
- Putro, H. F., Vulandari, R. T., & Saptomo, W. L. Y. 2020. Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan. Jurnal Teknologi Informasi Dan Komunikasi (Tikomsin), Vol. 8, No.2:19–24.
- Wardhana, et al. 2023. Implementasi Klasifikasi Naive Bayes dan Pemodelan Topik dengan Latent Dirichlet Allocation untuk Data Ulasan Video Game Lokal Pada Platform Steam. Emerging Statistics and Data Science Journal, Vol 1, No.3, Tahun 2023.